Working Papers Series 1

USR 3330
"Savoirs et Mondes Indiens"

HIERARCHICAL HIDDEN MARKOV STRUCTURE FOR DYNAMIC CORRELATIONS: THE HIERARCHICAL RSDC MODEL

Philippe Charlot* and Vêlayoudom Marimoutou*,†
GREQAM* and Institut Français de Pondichéry†

2011

Institut Français de pondichéry
Pondicherry

Centre de Sciences Humaines
New Delhi

Submitted to the Annals of Applied Statistics

HIERARCHICAL HIDDEN MARKOV STRUCTURE FOR DYNAMIC CORRELATIONS: THE HIERARCHICAL RSDC MODEL.

By Philippe Charlot[*]
and
Vêlayoudom Marimoutou[*,†]

GREQAM[*] and Institut Français de Pondichéry[†]

This paper presents a new multivariate GARCH model with time-varying conditional correlation structure, which is a special case of the Regime Switching Dynamic Correlation (RSDC) of [17]. This model, have named Hierarchical RSDC (HRSDC), has been built with the hierarchical generalization of the hidden Markov model introduced by [7]. This can be viewed graphically as a tree-structure with different types of states. The former are called production states, and they can emit observations, as in the class of Markov-Switching approach. The later are called "abstract" states. They can't emit observations but establish vertical and horizontal probabilities that define the dynamic of the hidden hierarchical structure. The main advantage of this approach, comparable to the classical Markov-Switching model that it improve the granularity of the regimes. Our model is also comparable to the new Double Smooth Transition Conditional Correlation GARCH model (DSTCC), a STAR approach for dynamic correlations proposed by [20]. The reason is that, under certain assumptions, the DSTCC and our model represent two classical competing approaches to modeling regime switching. We performed, Monte-Carlo simulations, and we applied the model to two empirical applications in studying the conditional correlations of selected stock returns. Results show that the HRSDC provides a good measure of the correlations, and possesses an interesting explanatory power.

1. Introduction. Since the seminal papers of [6] and [4], the study of multivariate GARCH with dynamic correlations has given rise to many extensions and developments. This growth in interest in the subject has been engendered by several empirical studies of stock market behaviours (e.g. [13–15]). They show that the hypothesis of constant correlations (the CCC model of [2]) is not realistic. However, the assumption of dynamic correlations is now widely accepted in the literature, the matter of the form

AMS 2000 subject classifications: Primary 62M10, 62H20; secondary 91B84, 62P20
Keywords and phrases: Multivariate GARCH, Dynamic correlations, Regime switching, Hierarchical Hidden Markov models

of the dynamic still remain an open question; see [1] or [19] for some recent surveys.

The purpose of this paper is to present a new multivariate GARCH model with dynamic correlations. We propose a regime switching model that is part of the Markov-Switching class. It is a special case of the RSDC model of [17]. This model is halfway between the CCC of [2] and the DCC of [6]. Correlations are constant within each regime, but vary from one regime to another and the transitions between different regimes are performed by a Markov chain. [18] have proposed a model with smooth transition between regimes for correlations (the STCC model), which can be seen, under certain assumptions, as a special case of the competitive STAR approach of Pelletier's model. The STCC requires the correlations to vary between two matrices of constant correlations. The transition between this two extreme matrices is governed by a conditional logistic function. Recently, they built an extension of this model, the Double-STCC (see [20]), in which conditional correlations vary across four matrices of constant correlations through two logistic functions. Our new model, the Hierarchical-RSDC (HRSDC), can be seen as a Markov-Switching version of the DSTCC. In this new model, correlations vary between four correlation matrices constant in time, but their transition from one matrix to another is determined by a hierarchical hidden Markov structure. The originality of this structure lies in its ability to establish a hierarchy between the hidden states in order to increase the granularity of the regime. This hierarchical hidden structure was first developed by [7] for handwriting recognition. In the present context, this specific structure will allow us to bring out a finer definition of regimes that do not conform to the classical Markov-Switching approach.

This paper is organized as follows. The HRSDC model is introduced in section 2.1.2. The HRSDC and the DCC is compared in section3.1. Section 3 presents results of Monte-Carlo simulations and two empirical applications and compare the HRSDC with the DSTCC and the DCC models. Section 4 presents concluding remarks and expose some directions for future research.

2. Hierarchical Hidden Markov Structure for Dynamic Correlations. The DSTCC model of [20] expresses the conditional correlation matrix as a combination of two extreme matrix that are themselves dynamics. By introducing another transition around the first one, the DSTCC works on two *primary* regimes, themselves built under four *secondary* regimes. This classification into *primart* and *secondary* regimes refers to the concept of sub-regime. The Markov-Switching approach of [17] doesn't consider the case of sub-regime to establish a hierarchy between states that define sub-

regimes, and show the nuances induced by the existence of sub-regimes in the dynamic of the correlations. As a first step, we will present the class of hierarchical hidden Markov models, the structure upon which our model has been built. Next we will explain our model.

2.1. *Structure.*

2.1.1. *The Hierarchical Hidden Markov Model.* The Hierarchical Hidden Markov Model (HHMM) was proposed by [7] in order to generalize the HMM model. The idea was to build a stochastic process with several levels by adopting a tree structure to obtain an interlacing of regimes. The hierarchy of the tree is built with *internal states*, which are abstract states (i.e. they do not produce observations). The internal states can lead to called *emitting states*, which produce observations. Internal and emitting states can also lead to a third type of state, which are called *exiting states*. The exiting states allow quitting from a level of the tree. Each internal states produces an sub-HMM, which can also be itself an HHMM. In this framework, the classical HMM is a special case of the HHMM with only one level. The main advantage of the HHMM compared to the HMM lies in its ability to improve the *granularity* of the regimes. The HHMM approach permits one to break up a time series into several types of regimes. For example, an HHMM with two levels has two types of regimes: *primary regimes* and *secondary regimes*. Certain combinations of the secondary regimes permits one to deduce the primary regimes. This increased granularity allows us to bring out nuances in regimes, something that is not possible with the simplified structure of an HMM.

Formally, an HHMM can be represented as the process $\{\mathcal{Y}_t, \mathcal{Q}_t\}_{t\in\mathbb{N}}$ where :

- $\{\mathcal{Y}_t\}_{t\in\mathbb{N}}$ is the process followed by the observations, which are supposed to be conditionally independent to the hidden states.
- $\{\mathcal{Q}_t\}_{t\in\mathbb{N}}$ is a homogeneous first order Markov chain. Each state of a HHMM q_i^d belongs to the set $\mathcal{Q} = \{\mathcal{S}, \mathcal{I}, \mathcal{E}\}$ where $\mathcal{S}$ is the set of emitting states, $\mathcal{I}$ the set of internal states and $\mathcal{E}$ the set of exiting states. The superscript d corresponds to the index Hierarchy (vertical location) in the tree, with $d \in \{1, ..., \mathcal{D}\}$ and subscript i is its horizontal location.

The tree structure, obtained by imposing an internal state at the root (level $d = 1$)[1]. This initial state may have several descendants which can be in-

[1]We will see later that this condition is not always necessary. The HRSDC model does

ternal or emitting states. These sub-internal states can themselves have descendants which can be internal or emitting states and so on. The transition from level d to level $d+1$ is provided by the probabilities of vertical transitions through an internal state. The return from level $d+1$ to level d is done with the exiting states, and corresponds to a probability of exiting. Internal and emitting states of the same sub-HMM of level d communicate with a transition matrix as in a classical HMM model. An internal state leading to another layer of internal/emitting states of a lower level is called a *parent* state. A parent state leads to *child* states. Finally, three probabilities govern the dynamics of the hidden structure:

- $\mathcal{A}_k^d = (a_k^d(i,j))$: is the matrix of horizontal transition of state i to state j where i and j are two states of the same sub-HMM of level d, i.e. $a_k^d(i,j) = \mathbb{P}[q_{t+1}^d = j | q_t^d = i]$, $q_{t+1}^d, q_t^d \in \{\mathcal{S}, \mathcal{I}\}$.
- $e_i^d = \mathbb{P}[q_{t+1}^d | q_t^{d+1}]$, with $q_t^{d+1}, q_{t+1}^d \in \{\mathcal{S}, \mathcal{I}\}$: vertical probability to leads from a child at level $d+1$ state to parent state at level d.
- $\pi_i^d = \mathbb{P}[q_{t+1}^{d+1} | q_t^d]$ with $q_t^{d+1}, q_{t+1}^d \in \{\mathcal{S}, \mathcal{I}\}$: vertical probability to lead from a parent state to its child state.

Each sub-model depending on parent k, must satisfy the conditions:

$$\sum_{j \in ch(k)} a_k^d(i,j) + e_i^d = 1 \text{ and } \sum_{i \in ch(k)} \pi_i^d = 1 \tag{2.1}$$

where $i, j \in ch(k)$ are two states with parents k. Figure 1 is an illustration of the basic structure of an HHMM. Following the notation of [21], , we shall denote by Q the size of the space state of each sub-model. Thus, the global configuration of a HHMM starting from the root up to the d^{th} level can be written as:

$$q^{(d)} = \overline{(q^1 q^2 ... q^d)} = \sum_{i=1}^{d} q^i Q^{d-i} \tag{2.2}$$

Thus, assuming that there is only one state at the root, all the parameters comprising an HHMM, satisfy the conditions:

$$\theta = (\cup_{d=2}^{D} \cup_{i=0}^{Q^{d-1}-1} \{\mathcal{A}_k^d, \pi_i^d, e_i^d\}) \cup (\cup_{i=0}^{Q^{D-1}} \{\phi_i\}) \tag{2.3}$$

where ϕ_i corresponds to the parameters of a probability distribution.

not possesses an internal state at the summit.

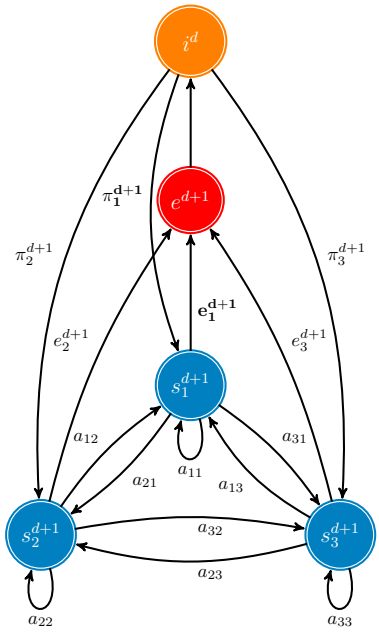

FIG 1. *Basic structure of a HHMM with three emitting states.*

2.1.2. *The Hierarchical RSDC model.* The objective of our model is to vary the correlations between two extreme major regimes, while allowing the existence of secondary regimes. As shown by [20], the correlation process is bounded by four states of constant correlations over time. The structure highlights two primary regimes, depending on abstract states i_1^1 and i_2^2. Each of these abstract states is connected with emitting states. Thus, the regime corresponding to i_1^1 is determinated by the emitting states s_1^2 and s_2^2; that of i_2^1 by s_3^2 et s_4^2. Figure 2 shows the hierarchical hidden structure of the Hierarchical RSDC model (HRSDC).

The hierarchical structure allows states to increase the granularity of the regimes. It establishes different types of regimes, which in our case are primaries and secondaries. The primary regimes correspond to the regimes obtained with a classical Markov Switching model. To ensure higher level of granularity, these primary regimes are build with sub-regimes, known as secondary regimes. The structure allows the secondaries's to capture nuances of dynamics that are finer than those of primaries's. The idea of granularity is illustrated on the figure 3.

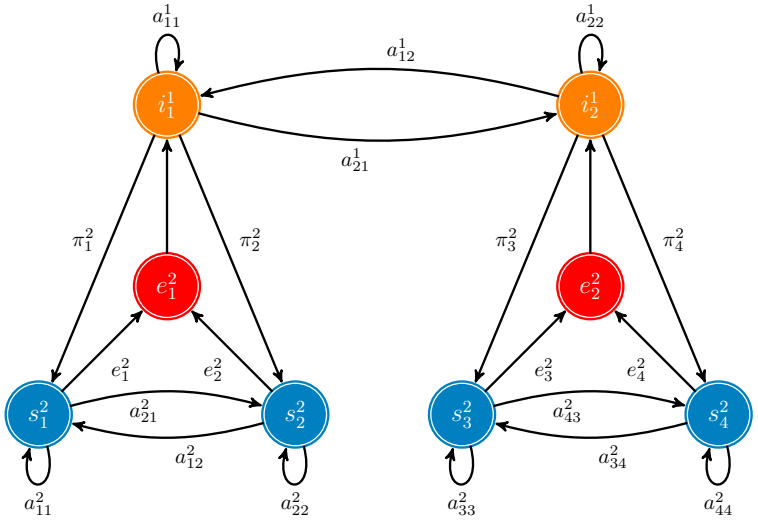

FIG 2. *Hierarchical Hidden structure of the HRSDC.*

The pair of emitting states defined by (s^2_1, s^2_2) forms a Markov-Switching model, and the same is true for (s^2_3, s^2_4). The link between these sub-models is provided by the abstract states i^1_1 and i^1_2. The model is then built on two sub-models with two emitting states each, for the transition matrices respectively[2] :

$$A^2_1 = \begin{bmatrix} a^2_{11} & a^2_{12} \\ a^2_{21} & a^2_{22} \end{bmatrix} \text{ and } A^2_2 = \begin{bmatrix} a^2_{33} & a^2_{34} \\ a^2_{43} & a^2_{44} \end{bmatrix}$$

and verified constraints:

$$\begin{cases} a^2_{11} + a^2_{21} + e^2_1 = 1 \\ a^2_{12} + a^2_{22} + e^2_2 = 1 \end{cases} \text{ and } \begin{cases} a^2_{33} + a^2_{43} + e^2_3 = 1 \\ a^2_{34} + a^2_{44} + e^2_4 = 1 \end{cases}$$

where e^2_i, $i = 1, ..., 4$ is the probability of exiting from a state of level two and go to a parent state at level one. The two sub-HMM, communicate via exiting states through abstract states i^1_1 and i^1_2. The activity involved in the transition from one to another of these abstract states is defined by the transition matrix:

$$A^1 = \begin{bmatrix} a^1_{11} & a^1_{21} \\ a^1_{12} & a^1_{22} \end{bmatrix}$$

[2]Recall that the probability that the state q which was in i at time $t-1$ is in j at time t is written $\mathbb{P}[q_t = j | q_{t-1} = i] = p_{ji}$.

which satisfies:

$$a_{11}^1 + a_{12}^1 = 1 \text{ and } a_{21}^1 + a_{22}^1 = 1$$

The parameters π_i^2, $i = 1, ..., 4$ represent the probability of moving from a parent state of the first level to one of its children at the second level. These probabilities must verify:

$$\pi_1^2 + \pi_2^2 = 1 \text{ and } \pi_3^2 + \pi_4^2 = 1$$

The specification for the four correlation matrices constants in time is that outlined by Pelletier. In fact, the only difference between this specification and the RSDC lies in the hierarchical hidden structure which led us to view the RSDC as a special case of the HRSDC with only one level. As in the RSDC model, the model can be estimated by either EM algorithm or iterative methods such as Gradient.

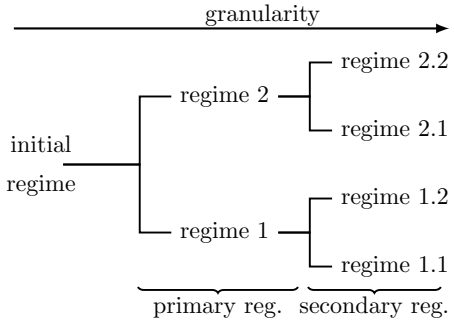

FIG 3. *The increase of the granularity in the HRSDC model.*

2.2. *Estimation.* Estimation under the HRSDC model is made using the multi-step estimation of [6] and [4]. This computationally attractive method splits up the log-likelihood as the sum of two parts : the volatility component and the correlation term. Estimation of the volatility part is done by maximizing the sum of the individual GARCH likelihoods. Estimation of the correlation part is trickier because of the abstract states involved. To estimate the second part, we follow the method of [21][3]. The main advantage of this method is the speed of implementation. Xie builds the transition

[3][21] succeeds in estimating the HHMM with the HMM's standard tools while respecting the vertical dynamic of the model. Xie's is built on a particular expression of the transition matrix. Instead of considering the whole dynamic of the model, the method breaks the transition matrix into several sub transition matrices for each hierarchical level. Each level of the tree is then linked with a transition matrix. This rewriting of the transition matrix allows one to estimate the correlations with a version of the Baum-Welch algorithm very similar to that used in the context of the standard HMM.

matrix by successive layers in order to have an expression of the likelihood of complete data that can be factorizable. In this framework, vertical transitions, which allows one to activate the child of a parent state, are given by:

$$\pi_q = \prod_{d=1}^{D} \pi^d_{q^d},\ q = 0, ..., Q^D - 1 \tag{2.4}$$

The transition matrix is rewritten by layers, each layer represents a level of the hierarchy of the tree:

$$\tilde{a}^d(q', q) = \prod_{i=d}^{D} e^i_{q'^i} \pi^i_{q^i} \cdot a(q'^d, q^d) \tag{2.5}$$

These probabilities are called *hypertransition probabilities*. Aggregation of these probabilities leads to the *hypertransition matrix* of our model:

$$\tilde{A}(q', q) = \sum_{i=1}^{D} \tilde{a}^i(q', q) \tag{2.6}$$

Finally, the hypertransition matrix has size $(S \times S)$ where $S = card(\mathcal{S})$. This formulation is attractive in the sense that it enables working with a transition matrix similar to the classical HMM or Markov-Switching model. The hypertransition matrix of our HRSDC model is then written as:

$$\tilde{A} = \begin{bmatrix} a^2_{11} + e^2_1 \pi^2_1 a^1_{11} & a^2_{21} + e^2_2 \pi^2_1 a^1_{11} & e^2_3 \pi^2_1 a^1_{12} & e^2_4 \pi^2_1 a^1_{12} \\ a^2_{21} + e^2_1 \pi^2_2 a^1_{11} & a^2_{22} + e^2_2 \pi^2_2 a^1_{11} & e^2_3 \pi^2_2 a^1_{12} & e^2_4 \pi^2_2 a^1_{12} \\ e^2_1 \pi^2_3 a^1_{21} & e^2_2 \pi^2_3 a^1_{21} & a^2_{33} + e^2_3 \pi^2_3 a^1_{22} & a^2_{43} + e^2_4 \pi^2_3 a^1_{22} \\ e^2_1 \pi^2_4 a^1_{21} & e^2_2 \pi^2_4 a^1_{21} & a^2_{34} + e^2_3 \pi^2_4 a^1_{22} & a^2_{44} + e^2_4 \pi^2_4 a^1_{22} \end{bmatrix} \tag{2.7}$$

The hidden hierarchical structure requires twenty parameters. However, in practice, only sixteen parameters will be needed in order to carry out estimations due to constraints of stochastic matrices. In the next sub-section, we propose two ways to estimate the correlations. The first is based on the EM algorithm. The second is done with Hamilton's filter.

2.2.1. *Estimation by EM algorithm.* To run EM algorithm, we need to write the quantity $\mathcal{Q}(\theta|\theta_k)$. With our hypertransition matrix, it is simply

written as:

$$
\begin{aligned}
(2.8)\quad \mathcal{Q}(\theta|\theta_p) &= \mathbb{E}_{\theta_p}[\log f(\epsilon_{1:T}, s_{1:T}; \theta)|\epsilon_{1:T}] \\
&= \sum_{t=1}^{T}\sum_{i=1}^{S} \log(f(\epsilon_t;\phi_i))\mathbb{P}_{\theta_p}[s_t = i|\epsilon_{1:T}] \\
&\quad + \sum_{t=1}^{T-1}\sum_{d=1}^{D}\sum_{i=1}^{S}\sum_{j=1}^{S} \log(\tilde{a}_{ij}^{d})\mathbb{P}_{\theta_p}[s_{t+1} = j, s_t = i|\epsilon_{1:T}] \\
&\quad + \sum_{i=1}^{S} \log(\nu_i)\mathbb{P}_{\theta_p}[s_1 = i|\epsilon_{1:T}]
\end{aligned}
$$

To simplify later formula 2.8, we define:

$$
\begin{aligned}
(2.9)\qquad & \gamma_t^p(q) \overset{def}{=} \mathbb{P}_{\theta_p}[q_t = q|\epsilon_{1:T}] \\
& \xi_t^p(q', q, d) \overset{def}{=} \mathbb{P}_{\theta_p}[q_t = q', q_{t+1} = q, e_t = d|\epsilon_{1:T}]
\end{aligned}
$$

Re-estimation formulas are obtained by maximizing the expected value of the complete-data log-likelihood $\mathcal{Q}(\theta|\theta_p)$. We refer to the work of [21] for a complete explanation of the procedure, and directly give the expression of the re-estimation formulas. Adopting [21]'s notations for indexing the states and writing $q = \overline{(rir')}$ and $q' = \overline{(rir'')}$, two states configurations which are identical up to the d^{th} level and have the same $(d-1)^{th}$ parent r such that $r = q^{1:d-1} = q'^{1:d-1}$ and $r' = q_t^{d+1:D}$ and $r'' = q_{t+1}^{d+1:D}$, two state configurations of a level below d, for each level $d \in \{1, ..., \mathcal{D}\}$, the re-estimation variables at iteration $p+1$ can be displayed as:

$$
\begin{aligned}
(2.10)\qquad & \bullet\ \hat{\pi}_r^d(j) = \frac{\sum_{t=1}^{T-1}\sum_{r'}\sum_{r''}\sum_{i}\xi_t^p(\overline{(rir')}, \overline{(rjr'')}, d)}{\sum_{t=1}^{T-1}\sum_{r'}\sum_{r''}\sum_{j}\sum_{i}\xi_t^p(\overline{(rir')}, \overline{(rjr'')}, d)} \\
& \bullet\ \hat{e}_r^d(i) = \frac{\sum_{t=1}^{T-1}\sum_{r'}\sum_{r''}\sum_{q'}\sum_{d'\leq d}\xi_t^p(\overline{(rir')}, q', d')}{\sum_{t=1}^{T-1}\sum_{r'}\sum_{r''}\gamma_t^p((rir'))} \\
& \bullet\ \hat{a}_r^d(i,j) = \frac{\sum_{t=1}^{T-1}\sum_{r'}\sum_{r''}\xi_t^p(\overline{(rir')}, \overline{(rjr'')}, d)}{\sum_{t=1}^{T-1}\sum_{r'}\sum_{r''}\sum_{j}\xi_t^p(\overline{(rir')}, \overline{(rjr'')}, d)} \times (1 - \hat{e}_r^d(i)) \\
& \bullet\ \hat{R}_q = \frac{\sum_{t=1}^{T}\epsilon_t \cdot \gamma_t^p(q) \cdot \epsilon_t'}{\sum_{t=1}^{T}\gamma_t^p(q)}
\end{aligned}
$$

We simply use as stopping rule the conventional difference $\|\hat{\theta}^{p+1} - \hat{\theta}^p\| \leq 10^{-6}$.

2.2.2. *Estimation by Gradient methods.* Using Gradient methods to estimate an HMM is made possible with the so-called Hamilton's filter (see [9–11]). This iterative filter allows one to make inference in the state of the unobserved Markov chain. With Hamilton's notations, let $\hat{\xi}_{t|t}$ be a vector of size $(N \times 1)$ which elements $\xi_{jt} = \Pr[s_t = j|\mathcal{F}_{t-1}, \theta]$, $j = 1, ..., N$, are the conditional probabilities to be in regime j, given the information set $\mathcal{F}_{t-1}$ at time $t-1$ and $\tilde{A}$ our hypertransition matrix of size $(N \times N)$ (see equation 2.6). Then one obtain a vector η_t of size $(N \times 1)$, of the elements are the densities under the N regimes, i.e. $f(\epsilon_t|s_t = j, \mathcal{F}_{t-1}; \theta)$, $j = 1, ..., N$. Then, Hamilton's filter gives the following filtered probabilities:

$$\hat{\xi}_{t|t} = \frac{(\hat{\xi}_{t|t-1} \odot \eta_t)}{\mathbf{1}'(\hat{\xi}_{t|t-1} \odot \eta_t)} \tag{2.11}$$

with forecasts calculated by:

$$\hat{\xi}_{t|t+1} = \tilde{A} \times \hat{\xi}_{t|t} \tag{2.12}$$

The inference for each date t is then found by iterating equations 2.11 and 2.12.

2.2.3. *Smoothed probabilities.* Hamilton's filter allows one to make inferences about the state of the Markov chain at time t conditional on the information set up to time t. [12] has developed one filter in order to make inferences with the whole information set. Instead of computing $\xi_{t|t}$, it permits to compute $\xi_{t|T}$ with $t < T$. Kim's filter is computed with our hypertransition as in the classical case:

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} \odot \{\tilde{A}'[\hat{\xi}_{t+1|T} \oslash \hat{\xi}_{t+1|t}]\} \tag{2.13}$$

where $\oslash$ denotes element-by-element division.

3. Applications. This section contains Monte-Carlo experiments and two applications based on real dada. The first real database is that used by [3] and contains daily data from S&P500 and 10-year bond futures. The second real database application is performed with that of [17]. It contains the exchange rates at the close of four week-days of the Pound, the Deutschmark, the Yen, and the Swiss-Franc against the US dollar.

3.1. *Simulated data.* In this sub-section, we compare the correlation estimates of our HRSDC model and the DCC_{ES} of [6] in a setting where the true correlation structure is known. For simplicity, it is done in a bivariate

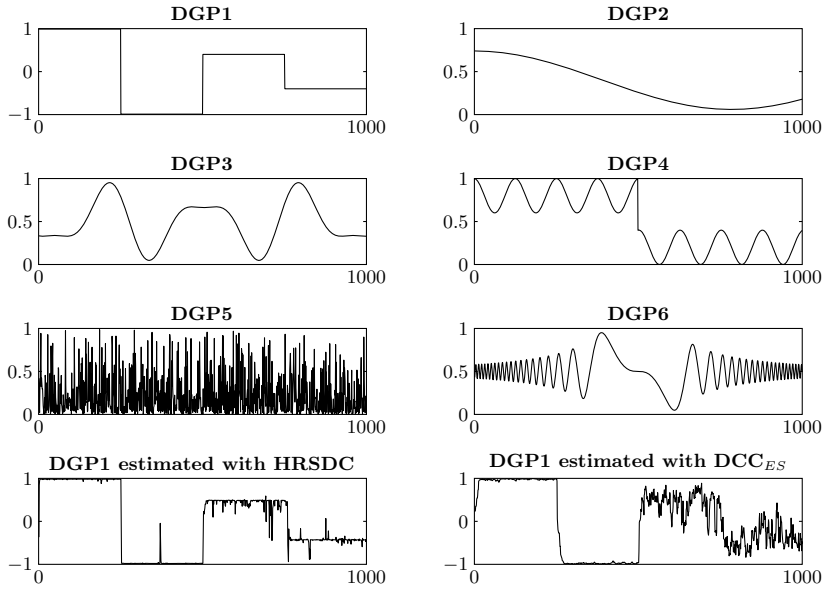

FIG 4. *Simulated Correlation Processes. The last two graphs are an example of the estimated correlations of the two models for DGP1.*

framework. We simulate six DGP. The first one has the following variance equation:

$$h_{1,t} = 0.01 + 0.05r_{1,t-1} + 0.85h_{1,t-1}$$
$$h_{1,t} = 0.12 + 0.1r_{1,t-1} + 0.85h_{1,t-1}$$

and the others (DGP2...DGP6):

$$h_{1,t} = 0.01 + 0.04r_{1,t-1} + 0.95h_{1,t-1}$$
$$h_{1,t} = 0.01 + 0.2r_{1,t-1} + 0.5h_{1,t-1}$$

For DGP1, the correlations process follows a TAR-CCC model with constant correlations. The DGP2...DGP6 are built with dynamic processes. The different DGP are labelled as:

- DGP1: $R_{12,t} = \begin{cases} 0.99 \text{ if } t \in [1;250] \\ -0.99 \text{ if } t \in [251;500] \\ 0.4 \text{ if } t \in [501;750] \\ -0.4 \text{ if } t \in [751;1000] \end{cases}$
- DGP2: $R_{12,t} = 0.4 + 0.34\cos(t/250)$
- DGP3: $R_{12,t} = 0.5 + \cos(d/s) - (1/3)\cos(3d) + (1/7)\cos(5d)$, $d = (t-50)/145$, $s = 35$

	DGP1	DGP2	DGP3	DGP4	DGP5	DGP6
MAE$_1$						
DCC	.1331	.5125	.3774	.4237	.5298	.2764
HRSDC	.0562	.5521	.3836	.4642	.5386	.2831
MAE$_2$						
DCC	.0553	.2901	.1700	.2361	.2824	.0778
HRSDC	.0235	.3408	.1660	.2659	.3008	.0835
DQ 5% *(strategy EW)*						
DCC	.7594	.5884	.6499	.9299	.5784	.9586
HRSDC	.5966	.6985	.7814	.9699	.4263	.9800
DQ 1% *(strategy EW)*						
DCC	.3359	.9944	.9970	$1e^{-4}$	.0217	.9925
HRSDC	.4382	.9733	.9995	$2e^{-4}$	.0215	.9929
DQ 5% *(strategy LS)*						
DCC	.7969	.1616	.5136	.9532	.7284	.7068
HRSDC	.8731	.0912	.3509	.9216	.6574	.7321
DQ 1% *(strategy LS)*						
DCC	.4294	$4.3e^{-5}$	.9986	.9996	.01	.0523
HRSDC	.3896	.2685	.9959	.9994	.01	.0136

TABLE 1
Performance measures results.

- DGP4: $R_{12,t} = \begin{cases} .8 + .2\cos(t/20) \text{ if } t \in [1;500] \\ 0.2 + .2\cos(t/20) \text{ if } t \in [501;1000] \end{cases}$
- DGP5: $R_{12,t} = 0.99 - \frac{1.98}{1+\exp(0.5\max(\epsilon^2_{1,t-1},\epsilon^2_{2,t-1}))}$
- DGP6: $R_{12,t} = 0.5 + \sin(s^3)/(1+\sqrt{|s^3|})$, $s = 5 - t/100$

The simulated correlations are voluntarily pathological in order to test the accuracy of the correlation estimates corresponding to a very volatile/distress periods of the financial markets. The DGP2, DGP4 and DGP6 are each built with sinusoidal functions to create correlations with different regimes or sub-regimes. The DGP5 was used by [16], and corresponds of the stylized fact pointed out by Longin and [15] that correlations among assets tend to increase during volatile periods.

The performance measures we use are very similar to those used by [4]. We first calculate two versions of a very classical loss function, which are computed as follows:

$$MAE_1 = \frac{1}{T}\sum_{t=1}^{T}\left|\hat{R}_t - R_t\right|$$
$$MAE_2 = \frac{1}{T}\sum_{t=1}^{T}(\hat{R}_t - R_t)^2$$

For the second type of measure, we follow the methodology of [4] by considering the loss function of the Value-at-Risk (VaR). Recall that for a portfolio with a share w invested in the first asset and $(1-w)$ in the second, the VaR assuming normality can be computed as:

$$VaR_t^{\alpha} = \Phi_t^{-1}(\alpha)\sqrt{(w^2\hat{H}_{11,t} + (1-w)^2\hat{H}_{22,t} + 2w(1-w)\hat{R}_{12,t}\sqrt{\hat{H}_{11,t}\hat{H}_{22,t}})}$$

The loss function of the VaR is then defined by:

$$hit_t = \mathbb{1}_{\{wr_{1,t}+(1-w)r_{2,t}<-VaR_t\}} - \alpha$$

We then use the *in-sample Dynamic Quantile* (DQ) test introduced by [5]. This consists of checking if all the regression coefficients of the violation process hit_t with its lagged values and others exogenous variables are equal to zero. To perform this F test, we use as explanatory variables ,five lags hit_t and the current value of the VaR. We test two numbers of rejections (1% and 5%) and two portfolios: an equal-weighted ($w = 0.5$, strategy EW) and long-short (1 and -1, strategy LS). Table 1 presents the results of the performance measures.

The results show that the HRSDC performs better than the DCC when correlations are constant within a sample. This result is consistent with what could be expected since the HRSDC varied correlations between several constant correlations matrices over time. Once the simulated correlations are no longer constant within a specific sample, the DCC model has better MAE, a result which, once again, seems normal. However, differences in the MAE between the DCC and the HRSDC are comparatively small. The dynamic process from a combination of regimes of the HRSDC does not too appear to fare too badly in comparison to the autoregressive motion of the DCC.

The results of the DQ test are more mixed, and appear highly dependents on the strategy chosen. The DCC is the best for the EW strategy, while the HRSDC dominates on the LS strategy. A simple addition of the best values show that the HRSDC seems preferable to DCC. Nevertheless, it would be unfair to conclude to the dominance of HRSDC under the DCC. It is just better to point out that the HRSDC remains credible face to the DCC.

Finally, the HRSDC remains a tool of great efficiency when correlations are constant within a given sample (DGP1), but becomes less effective when that is no longer the case (DGP2...DGP6). It seems that the autoregressive activity of the DCC remains very efficient when correlations oscillate

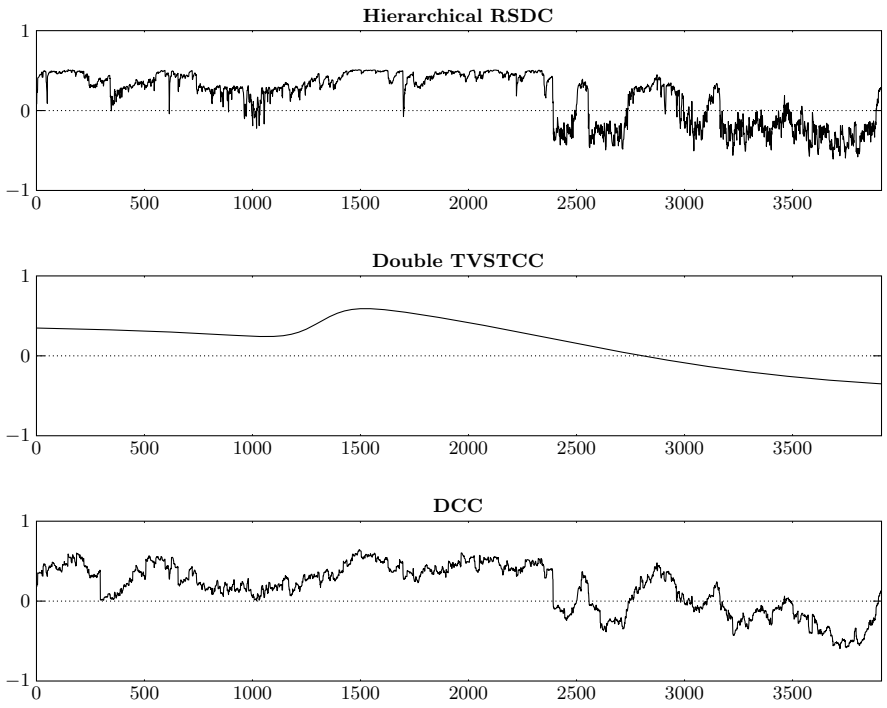

FIG 5. *Estimated correlations of the HRSDC, the DCC_{ES}, and the DSTCC with the sample of Colacito and Engle (2006).*

around a trend value. However, the appeal of the HRSDC lies in its explanatory power. Apart from providing a simple measure of correlations, its structure allows one to bring out the components of their overall variability by increasing the granularity of regimes.

3.2. *Correlations between S&P500 index futures and 10-year bond futures.* The first application with a real database is based on the bivariate sample of [3]. It contains daily returns of S&P500 index futures and 10-year bond futures from January 1990 to August 2003. These data are also use by [20] in demonstrating to apply their DSTCC model. The individuals volatilities are obtained by running a GARCH(1,1) model. We later obtain the correlations by running three different models in order to make comparisons. These models are the HRSDC, the DCC_{ES} and the DSTCC. For this sample, the correlations of the DCC_{ES} and DSTCC are estimated using Gradient methods while the correlations of the HRSDC are computed with EM algorithm. Figure 5 shows the estimated correlations for the three models. In our application, transition variables for the DSTCC are defined as calendar

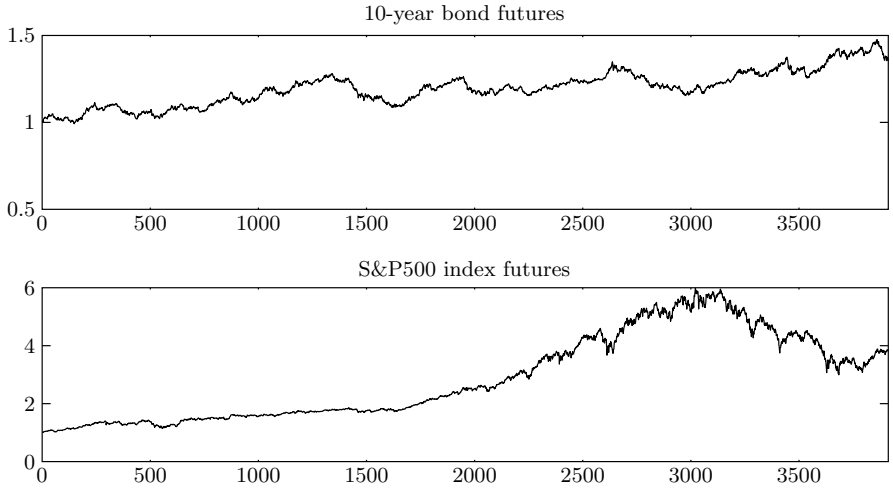

FIG 6. *S&P500 index futures and 10-year bond futures.*

time.

Results of the estimated parameters of the HRSDC are presented in Appendix A and the smoothed probabilities corresponding to the secondary regimes are in Figure 7. The model has correctly identified two sub-regimes with positive correlations and two sub-regimes with negative correlations. Adding the smoothed probabilities of each pair of sub-regimes gives the smoothed probabilities that we would have had with a Markov-Switching model with two regimes. This is shown in Figure 8. The results for the DSTCC (with the calendar time specification) show that it does not accurately capture the *chronological* changes in a regime. Correlation estimates have the form of a curve. Rather than a local measure, this model serve to indicate a trend. For example, in the DCC and the HRSDC models, correlations go to a negative correlation regime around the 2400^{th} observation whereas they are still positives in the DSTCC. This is due to the fact that the DSTCC has only two transition functions. A solution to this problem could be to introduce more transition functions as in Amado and Teräsvirta (2008). Even though the DCC and HRSDC correlations seem similar, in the case study, the HRSDC has the advantage that it can explain the nuances in the dynamics through the decomposition in the sub-regimes. An increase The increasing in granularity ends up in a finer definition of the transition.

3.3. *Correlations between exchange rate data.* In this second application, we apply our HRSDC model to the sample used by [17]. This series are

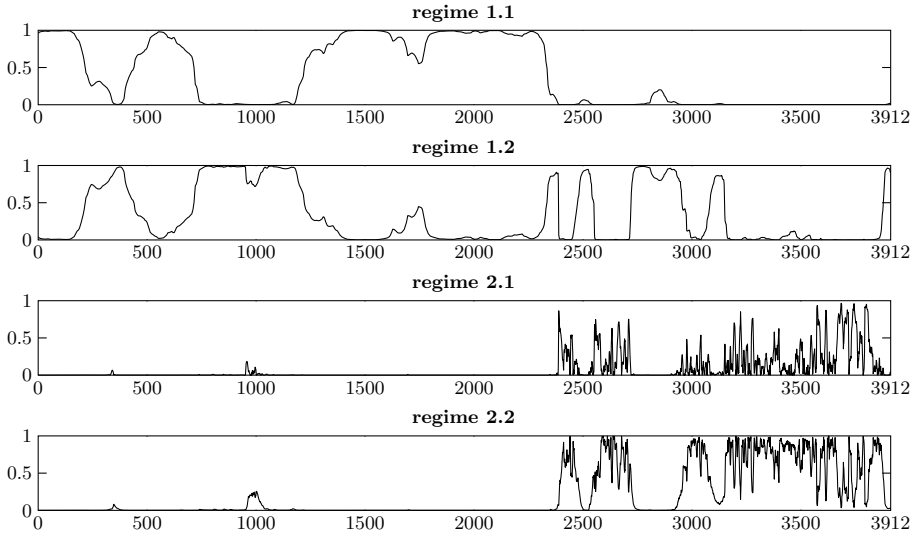

FIG 7. *Smoothed probabilities of the HRSDC. Regimes 1.1 and 1.2 correspond to the emitting states* $\{s_1^2, s_2^2\}$*, and regimes 2.1 and 2.2 to* $\{s_3^2, s_4^2\}$*.*

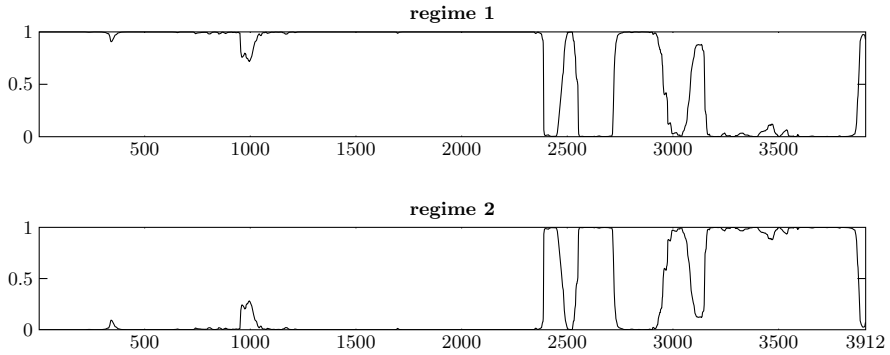

FIG 8. *Smooth probabilities of the regimes defined by the internal states* i_1^1 *and* i_2^1 *(first application).*

exchange rate data and are plotted as in Figure 9. As before, the motion of the standard deviations is obtained with a GARCH(1,1) model. For this case study, only DSTCC and HRSDC models have been considered. We use the two-step maximum likelihood estimation using Gradient methods for both models. Figure 10 shows the estimated correlations plotted for the two models. Parameters estimated for the HRSDC are shown in the Table 3 (see Appendix A). Smooth probabilities for the HRSDC can be

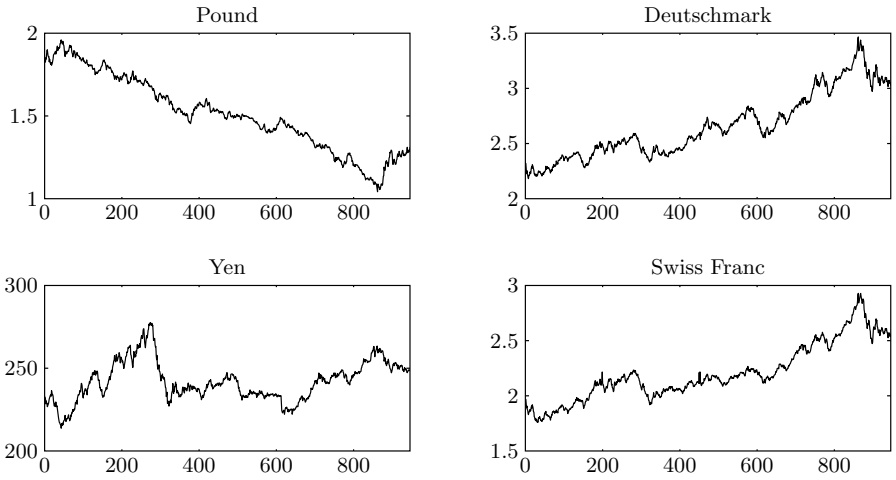

FIG 9. *Exchange rate database.*

seen in Figure 11. As we see from Figure 10, as in the previous application, the DSTCC (with its calendar time specification, i.e. $m_{1t} = m_{2t} = t/T$) allows one at best to perceive a trend in the evolution of correlations in the sample. The reason for this limitation is often that, with only two transition functions, the DSTCC cannot capture more than three regimes. A direct consequence of this shortcoming is that the DSTCC fails to capture the time specific extreme regimes, as it demonstrated in example on the Swiss Fr/Deutschmark correlations. These are indeed marked by two peaks of correlations close to zero as they oscillate the rest of the time around a value about 0.8.

In this sample, the HRSDC clearly identifies four sub-regimes. Recall that the first primary regime associated with the internal state i_1^1 is a combination of secondary regimes depending of the two emitting states s_1^2 and s_2^2. Increasing the granularity provided by the hidden hierarchical structure allows one to highlight the existence of a sub-regime linked to s_2^2. This sub-regime occurs only very rarely within the sample, around the 450^{th} observation. But it permits one to capture an extreme and time specific behaviour of the correlations process. This element of the global activity of correlations would be go unnoticed with a classical Markov-Switching model (because that model is too limited and not significant enough in relation to the size of the sample).

This application with a sample of four series brings to light a problem in the correlations specification. As we have said before, estimated parameters

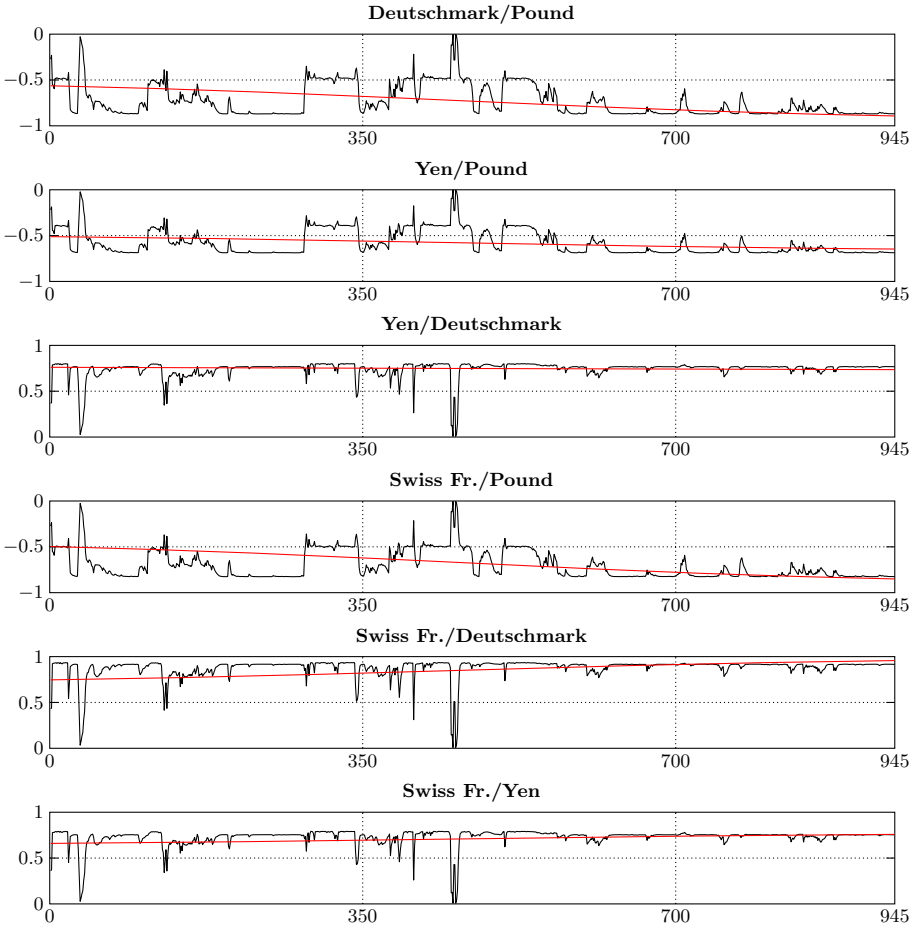

FIG 10. *Estimated correlations for the HRSDC (in black) and the DSTCC (in red).*

have been obtained using Gradient methods to estimate the HRSDC. This choice is not fortuitous. In fact, our experiments shows that the non constraint specification meets with difficulties in returning the maximum of the objective function with the use of EM algorithm. The cause of the problem is to be found in the constraint of the Choleski representation introduced in the correlation matrix. Without this transformation, the derivatives of the objective function produce a form of the correlation matrix with no guarantees to have one in the diagonal and PSD. The transformed Choleski representation without one in the diagonal greatly disrupts the convergence of the EM algorithm. That's why it is preferable to use the constraint specification which can be estimated with the Gradient method using Hamilton's

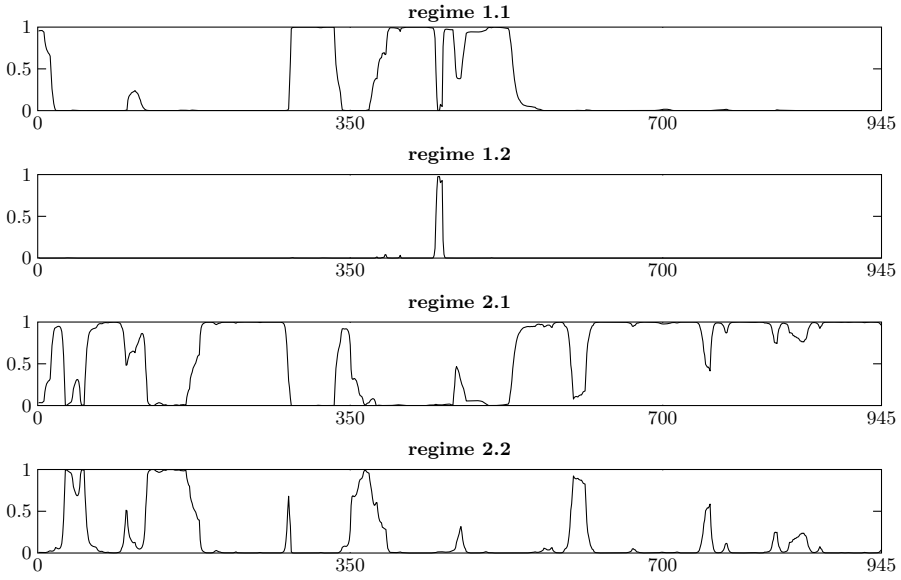

FIG 11. *Smoothed probabilities for the HRSDC (exchange rate data).*

filter. In this example, despite a huge number of iterations, the EM algorithm has not converged whereas Gradient methods rapidly reach the optimum of the objective function. Nevertheless, even the constraint specification would need numerous parameters, and the estimation of large correlation matrices could turn out to be cumbersome.

4. Conclusion. In this paper, we have presented a new multivariate GARCH with dynamic correlations. This further extension, called Hierarchical RSDC (HRSDC), can be view as a special case of the RSDC model of [17]. The HRSDC is a Markov-Switching class model with a correlation process bounded by four correlation matrices constants over time. The innovation of our model is that it is built on a hierarchical hidden structure introduced by [7].

The main advantage of this hidden tree-like structure lies in its ability to increase the granularity of the regime. It permits one to define different types of regime; in our case, primary and secondary regime. Applied to correlation modeling, the HRSDC allows one to capture finer nuances than is possible with the classical Markov-Switching approach. Monte Carlo experiments and applications on real data show that this approach could improve understanding of the randomness of correlations. The application of the HRSDC to estimate the correlations between S&P500 index futures

and 10-year bond futures, between exchange rate data has brought to light the existence of sub-regimes which other regime switching models have so far unable to do.

While the results in this paper show that the HRSDC possesses good explanatory powers, one must be aware that it has limitations as well. The first concerns selection of models. Our model has been buit from a symmetric hidden tree, with two primary regimes, each of them with which one sub-HMM. This represents a very simple structure. It is possible to build an asymmetric tree with severals levels of depth. However, finding the best hierarchy, i.e. split/merge/swap levels, could be a present problem. [21] use the Reversible-Jump MCMC (RJMCMC) of [8] in the selection of this models. While the approach may appear attractive, one will have to contend with the questions of size that it entails. The second problem relates to the specification of the correlation matrix. The specification of Pelletier is not suitable in the modeling of large correlation matrices. As the hierarchical hidden structure model is a plug-in method, it could be interesting to find a specification for large correlation matrices. The tasks awaits future research.

APPENDIX A: ESTIMATED PARAMETERS OF THE APPLICATIONS

Table 3 shows the parameters estimations of the correlations process for the database of exchange rate data; table 2 for the database of Engle and Colacito.

correlation matrix:

$$R_1 = \begin{bmatrix} 1 & \\ 0,5189 & 1 \end{bmatrix} \qquad R_2 = \begin{bmatrix} 1 & \\ 0.2924 & 1 \end{bmatrix}$$

$$R_3 = \begin{bmatrix} 1 & \\ -0.1853 & 1 \end{bmatrix} \qquad R_4 = \begin{bmatrix} 1 & \\ -0.1597 & 1 \end{bmatrix}$$

transition probabilities

$$A^1 = \begin{bmatrix} 0.9797 & 0.9365 \\ 0.0203 & 0.0635 \end{bmatrix}$$

$$A_1^2 = \begin{bmatrix} 0.9919 & 1.4e^{-4} \\ 0.0055 & 0.7876 \end{bmatrix} \quad A_1^2 = \begin{bmatrix} 0.8139 & 0.0618 \\ 0.1853 & 0.9299 \end{bmatrix}$$

$$e_1^2 = 5.85e-4,\ e_2^2 = 0.2123,\ e_3^2 = 7.7e-4,\ e_4^2 = 0.0084$$

$$\pi_1^2 = 0.0067,\ \pi_2^2 = 0.9933,\ \pi_3^2 = 0.9933,\ \pi_4^2 = 0.0067$$

TABLE 2

Estimated parameters for the correlations of the second real data application.

correlation matrix:

$$R_1=\begin{bmatrix} 1 & & & \\ -0,4860 & 1 & & \\ -0,3943 & 0,8125 & 1 & \\ -0,5004 & 0,9488 & 0,8032 & 1 \end{bmatrix} \quad R_2=\begin{bmatrix} 1 & & & \\ -4e-5 & 1 & & \\ -3e-5 & 8e-05 & 1 & \\ -5e-5 & 9e-5 & 8e-5 & 1 \end{bmatrix}$$

$$R_3=\begin{bmatrix} 1 & & & \\ -0,8746 & 1 & & \\ -0,6912 & 0,7720 & 1 & \\ -0,8299 & 0,9214 & 0,7587 & 1 \end{bmatrix} \quad R_4=\begin{bmatrix} 1 & & & \\ -0,7337 & 1 & & \\ -0,5798 & 0,6476 & 1 & \\ -0,6962 & 0,7730 & 0,6365 & 1 \end{bmatrix}$$

transition probabilities

$$A^1=\begin{bmatrix} 0,9933 & 0,4167 \\ 0,0066 & 0,5832 \end{bmatrix}$$

$$A_1^2=\begin{bmatrix} 0,6711 & 0,0084 \\ 0,0115 & 0,9915 \end{bmatrix} \quad A_1^2=\begin{bmatrix} 0,9835 & 0,0002 \\ 0,0157 & 0,9571 \end{bmatrix}$$

$e_1^2 = 0.3172$, $e_2^2 = 8.21e-05$, $e_3^2 = 7.37e-04$, $e_4^2 = 0.0425$

$\pi_1^2 = 0.9933$, $\pi_2^2 = 0.0066$, $\pi_3^2 = 0.9933$, $\pi_4^2 = 0.0067$

TABLE 3

Estimated parameters for the correlations of the second real data application.

REFERENCES

[1] BAUWENS, L., LAURENT, S. and ROMBOUTS, J. V. (2006). Multivariate GARCH Models : a Survey. *Journal of Applied Econometrics* **21** 79–109.

[2] BOLLERSLEV, T. (1990). Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH model. *Review of Economics and Statistics* **72** 498–505.

[3] COLACITO, R. and ENGLE, R. F. (2006). Testing and Valuing Dynamic Correlations for Asset Allocation. *Journal of Business and Economic Statistics* **24** 238–253.

[4] ENGLE, R. F. (2002). Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models. *Journal of Business and Economic Statistics* **20** 339–350.

[5] ENGLE, R. F. and MANGANELLI, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics* **22** 367–381.

[6] ENGLE, R. F. and SHEPPARD, K. (2001). Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH Working Paper report No. 2001–15, UCSD.

[7] FINE, S., SINGER, Y. and TISHBY, N. (1998). The Hierarchical Hidden Markov Model : Analysis and Applications. *Machine Learning* **32** 41–62.

[8] GREEN, P. (1995). Reversible Jump MCMC Computation and Bayesian Model Determination. *Biometrika* **82** 711–732.

[9] HAMILTON, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* **57** 357–384.

[10] HAMILTON, J. D. (1990). Analysis of Time Series Subject to Changes in Regime. *Journal of Econometrics* **45** 39–70.

[11] HAMILTON, J. D. (1994). *Time Series Analysis*. Princeton University Press.

[12] KIM, C. J. (1994). Dynamic Linear Models with Markov-Switching. *Journal of*

Econometrics **60** 1–22.
[13] Longin, F. and Solnik, B. (1995). Is the Correlation in International Equity Returns Constant : 1960–1990 ? *Journal of International Money and Finance* **14** 3–26.
[14] Longin, F. and Solnik, B. (1996). The Asymptotic Distribution of Extreme Stock Market Returns. *Journal of Business* **69** 383–408.
[15] Longin, F. and Solnik, B. (2001). Extreme Correlation of International Equity Markets. *Journal of Finance* **56** 651–678.
[16] Long, X. and Ullah, A. (2005). Nonparametric and Semiparametric Multivariate GARCH Model.
[17] Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics* **131** 445–473.
[18] Silvennoinen, A. and Teräsvirta, T. (2005). Multivariate Autoregressive Conditional Heteroskedasticity with Smooth Transitions in Conditional Correlations Working Paper Series in Economics and Finance report No. 0577, Stockholm School of Economics.
[19] Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate GARCH models. In *Handbook of Financial Time Series* (T. G. Andersen, R. A. Davis, J. P. Kreiss and T. Mikosch, eds.) Springer, New York.
[20] Silvennoinen, A. and Tersvirta, T. (2009). Modeling Multivariate Autoregressive Conditional Heteroskedasticity with the Double Smooth Transition Conditional Correlation GARCH Model. *Journal of Financial Econometrics* **7** 373-411.
[21] Xie, L. (2005). Unsupervised Pattern Discovery for Multimedia Sequences PhD Thesis report, Columbia University.

GREQAM
Centre de la Vieille Charité
2, rue de la Charité
13236 Marseille cedex 02
France
E-mail: ph.charlot@gmail.com
URL: http://greqam.univ-mrs.fr/

Institut Français de Pondichéry
11, Saint Louis Street
Pondicherry - 605 001
India
E-mail: v.marimoutou@ifpindia.org
URL: http://www.ifpindia.org/